



The Effects of Social Ties on Coordination: Conceptual Foundations for an Empirical Analysis

Giuseppe Attanasi, Astrid Hopfensitz, Emiliano Lorini, Frédéric Moisan

► To cite this version:

Giuseppe Attanasi, Astrid Hopfensitz, Emiliano Lorini, Frédéric Moisan. The Effects of Social Ties on Coordination: Conceptual Foundations for an Empirical Analysis. *Phenomenology and the Cognitive Sciences*, 2014, 13 (1), pp.47-73. 10.1007/s11097-013-9320-4 . hal-01123762

HAL Id: hal-01123762

<https://hal.science/hal-01123762>

Submitted on 5 Mar 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Open Archive TOULOUSE Archive Ouverte (OATAO)

OATAO is an open access repository that collects the work of Toulouse researchers and makes it freely available over the web where possible.

This is an author-deposited version published in : <http://oatao.univ-toulouse.fr/>
Eprints ID : 12584

To link to this article : DOI :10.1007/s11097-013-9320-4
URL : <http://dx.doi.org/10.1007/s11097-013-9320-4>

To cite this version : Attanasi, Giuseppe and Hopfensitz, Astrid and Lorini, Emiliano and Moisan, Frédéric *[The Effects of Social Ties on Coordination: Conceptual Foundations for an Empirical Analysis](#)*. (2014) Phenomenology and the Cognitive Sciences, vol. 13 (n° 1). pp. 47-73. ISSN 1568-7759

Any correspondence concerning this service should be sent to the repository administrator: staff-oatao@listes-diff.inp-toulouse.fr

The effects of social ties on coordination: conceptual foundations for an empirical analysis

Giuseppe Attanasi · Astrid Hopfensitz ·
Emiliano Lorini · Frédéric Moisan

Abstract This paper investigates the influence that social ties can have on behavior. After defining the concept of social ties that we consider, we introduce an original model of social ties. The impact of such ties on social preferences is studied in a coordination game with outside option. We provide a detailed game theoretical analysis of this game while considering various types of players, i.e., self-interest maximizing, inequity averse, and fair agents. In addition to these approaches that require strategic reasoning in order to reach some equilibrium, we also present an alternative hypothesis that relies on the concept of team reasoning. After having discussed the differences between the latter and our model of social ties, we show how an experiment can be designed so as to discriminate among the models presented in the paper.

Keywords Social ties · Coordination · Game theory · Forward induction · Social preferences · Team reasoning

Funding through the ANR-2010 JCJC 1803 01 is gratefully acknowledged.

G. Attanasi
Bureau d'Économie Théorique et Appliquée (BETA), Université de Strasbourg, 61, Avenue de la
Forêt-Noire, 67000 Strasbourg, France
e-mail: attanasi@unistra.fr

A. Hopfensitz
Toulouse School of Economics, 21 Allée de Brienne, 31000 Toulouse, France
e-mail: Astrid.Hopfensitz@univ-tlse1.fr

E. Lorini
IRIT - CNRS, University of Toulouse, 118 route de Narbonne, 31062 Toulouse Cedex 9, France
e-mail: Emiliano.Lorini@irit.fr

F. Moisan (✉)
IRIT, University of Toulouse, 118 route de Narbonne, 31062 Toulouse Cedex 9, France
e-mail: moisan@irit.fr

Introduction

In classical economic theories, most models assume that agents are self-interested and maximize their own material payoffs. However, important experimental evidence from economics and psychology has shown some persistent deviation from such self-interested behavior in many strategic situations. These results suggest the need to incorporate social preferences into game theoretical models. Such preferences describe the fact that a given player not only considers his own material payoffs but also those of other players (Margolis 1982). The various social norms created by the cultural environment in which human beings live give some ideas of how such experimental data could be interpreted: fairness, inequity aversion, reciprocity, and social welfare maximization are concepts that behavioral economists are familiar with, and which have been shown to play an important role in interactive decision making (e.g., see Charness and Rabin 2002; Fehr and Schmidt 1999; Rabin 1993a).

In fact, various simple economic games, such as the trust game (Berg et al. 1995) and the ultimatum game (Güth et al. 1982), have been extensively studied in the past years because they illustrate well the weakness of traditional game theory and its assumption of individualistic rationality. Moreover, given the little complexity carried out in such games, the bounded rationality argument (Gigerenzer and Selten 2001) does not seem sufficient to justify observed behavior. Social preferences appear as a more realistic option because they allow to explain the resulting behaviors while still considering rational agents.

However, although many economic experimental studies (e.g., Berg et al. 1995; Güth et al. 1982) have shown that people genuinely exhibit other-regarding preferences when interacting with perfect strangers, one may wonder to what extent the existence of some social relationships between individuals may influence behavior. In this article, we will refer to such social relationships as “social ties.” Indeed the dynamic aspect of social preferences seems closely related to that of social ties: one may cooperate more with a friend than with a stranger, and doing so may eventually enforce the level of friendship. Yet, in spite of their obvious relevance to the study of human behavior, very little is known about the nature of social ties and their actual impact on social interactions.

Our attempt, through this paper, is to study the possible effects that positive social ties can have on human cooperation and coordination. Our main hypothesis is that such relationships can influence a player’s choice by modifying his preferences: an agent may choose to be fair conditionally on the relative closeness to his partner(s). In order to investigate these questions, we propose a theoretical analysis of a new kind of two-player game that allows us to disentangle predictions from theories based on self-interest, social preferences, and social ties. Furthermore, we demonstrate the need to introduce an alternative model to capture the concept of social ties as continuous variables. Indeed, while we claim that social ties strongly rely on group identification, we show that considering the concept of team reasoning is too limited to fill this purpose as it is built upon a binary interpretation of group identification (i.e., either one identifies with a group or not).

The rest of the article is organized as follows. “A definition of social ties” defines the concept of a social tie that we consider. Through “A theory of how to model social ties,” we introduce a formal model that embeds social ties and allows to specify

players' behavior in strategic interactions accordingly. As a means to evaluate our theory, we propose, in "[A social ties game](#)," a game that allows to measure the behavioral effects of social ties. We provide, in "[Equilibrium predictions with self-interested players](#)," a game theoretical analysis of this game by considering only self-interested agents. In "[Equilibrium predictions under models of social preferences](#)," we perform a similar analysis by considering other-regarding agents according to theories of social preferences. After analyzing this strategic scenario through our model of social ties in "[Equilibrium predictions under our model of social ties](#)," we propose an alternative interpretation of the same game by considering agents as team-directed reasoners through "[Players as team-directed reasoners](#)." "[Why team reasoning cannot express gradual social ties](#)" then illustrates the relevance of our model of ties by discussing the limited expressivity of team reasoning in the context of social ties. Finally, in "[Further hypotheses](#)," we evaluate various models in the context of social ties and we state some relevant behavioral hypotheses.

A definition of social ties

No formal definition of a social tie is provided either in the literature on social psychology or in the experimental economics literature focused on social preferences. Thus, given the vagueness and the ambiguity that the term may suggest, we begin by clarifying the concept that we consider.

First, we choose to restrict our study only to those ties that can be judged to be positive: examples include relationships between close friends, married couples, family relatives, classmates, etc. In contrast, negative ties may include relationships between people with different tastes, from different political orientations, with different religious beliefs, etc.

In order to specify the foundations of such social ties and the possible reasons for their emergence, let us consider the well-known concept of social identity from social psychology. According to social identity theory (Hogg 2002; Tajfel and Turner 1979), an individual's social identity is built upon a set of social features, each of which may refer to any type of salient characteristics that can be shared by individuals in a particular context. For example, a person may identify himself as a student of the University of Toulouse, a supporter of Barcelona's soccer team, a Democrat, a Catholic, etc.

According to various theories in social psychology (see, e.g., Abrams and Hogg 2006; Hogg 2000), the construction of an individual's social identity is determined by two complementary motivations. The first motivation is self-enhancement which is underpinned by one's individual need to promote self-esteem (as pointed out by Luhtanen and Crocker (1992), "Being a member of a social group is an important reflection of who I am"). Reduction of subjective uncertainty about one's perceptions, attitudes, feelings, behavior, and one's self-concept and place within the social world is the second motivation.

It can be reasonably assumed that people can give different degrees of importance to those social features defining their social identity, depending on the context: for example, while one's identification as a soccer player is more important than one's

identification as a student during a soccer game, the reverse may hold for the same individual during a math exam at the university.

Following this interpretation, our claim is that:

Statement 2.01: *A social tie between two individuals exists if and only if they share the same social features defining their social identities, and this is common belief among them.*

Note that the previous claim implies that a *social* tie is necessarily *bilateral* in the sense that if an individual *i* is tied with another individual *j*, then *j* is also tied with *i*. For example, an individual who believes to share the same political convictions with a given politician cannot induce a social tie as long as the latter does not also believe so (one could speak of the existence of a unilateral tie in this case, though it is not “social” according to the above statement).

Moreover, the previous statement simply characterizes the minimal condition for the existence of a social tie. As an illustration, one can consider the well-known Minimal Group Paradigm (MGP) (Tajfel 1970), which corresponds to an experimental methodology from social psychology that investigates the minimal conditions required for discrimination to occur between groups. Experiments using this approach (Tajfel et al. 1971) have revealed that arbitrary and virtually meaningless distinctions between groups (e.g., the color of their shirts) can trigger a tendency to cooperate more with individuals within one’s own group than with others. In this case, one should note that such meaningless social features satisfy the minimal condition for being considered as a social tie from the previous statement. However, in principle such social tie should be quite weak.

In this respect, it is worth mentioning that an important property of social ties lies in its quantitative aspect, that is, two individuals can be *more or less* socially tied with each other. To be more precise, we assume that a social tie between two individuals can be measured on a scale ranging from 0 to 1, where 0 and 1 respectively stand for the minimum and maximum strength for the tie.

This interpretation therefore suggests that the strength of a social tie can be determined by the *quantity and importance of shared social features*. One can indeed assume that sharing a high number of social features (defining one’s social identity) with high importance leads to a high social tie value. On the other hand, having conflicting social characteristics, or sharing a low number of features with high importance, or sharing a high number of features with low importance can lead to a lower tie value.

Moreover, another aspect that, we believe, influences the strength of a social tie between two individuals is the *quantity and quality of past interactions* between them. More precisely, given two individuals sharing a certain number of social features with a given importance, the strength of the tie between them is higher in the situation in which the two individuals had frequent meaningful interactions in the past than in the situation in which there were no previous meaningful interactions.¹

¹ With the term “meaningful” we mean that during past interactions, the two individuals had the occasion to know each other by exchanging ideas, opinions, sharing positive emotions (e.g., they mutually enjoyed playing tennis together), etc.

As a concrete example to illustrate the previous interpretation, one may consider the case of online dating systems on the internet. Those systems, which are clearly meant to build social ties between individuals (assuming an affective tie is a special case of a social tie), are based on the matching of social features that define their social identities. However, while one cannot deny the effectiveness of such systems (Hitsch et al. 2010), it is suggested in Frost et al. (2008) that some interaction between two individuals is also important as it can allow them to know each other more accurately. Indeed, by providing a way to obtain reliable information about one another, social interactions happen to be a relevant tool against possibly inaccurate stereotypes, which can often be considered as an unfortunate consequence of categorizing individuals into social groups, as implied by social identity theory.

The following points summarize our interpretation of social ties:

- The minimal criterion for the existence of a social tie between two individuals is for them to commonly believe that they share the same social features that define their social identities.
- A social tie between two individuals has a quantitative dimension which depends on the following two parameters:
 1. The quantity and importance of shared social features that define both individuals' social identities.
 2. The quantity and quality of past interactions between both individuals.

Following our interpretation, one might then argue that the situation described by the minimal group paradigm satisfies the minimal condition for the existence of a social tie, even though this tie has a relatively low degree of strength (the number of shared social features is one) and its importance might be considered to be reasonably low.

A theory of how to model social ties

In this section, we introduce a novel model that characterizes the agents' behavior in the presence of social ties. Our model of social ties shares features with both *team reasoning* and *social preferences* theories.

Similarly to team reasoning theories, our model is built on the concept of group identification, which is of high relevance when considering social ties. In fact, individuals that are socially connected may be expected to identify themselves with the same group, which may consequently lead them to choose actions as a member of this group. In “[Players as team-directed reasoners](#),” we discuss various theories of team reasoning, at the same time indicating which properties are in common with our approach to social ties. In “[Why team reasoning cannot express gradual social ties](#),” we underline the inadequacy of such theories to interpret in full the effects of social ties, thereby claiming for a novel approach—the one introduced in this section—able to capture specific key features of social ties left aside by team reasoning.

Similarly to theories of social preferences, our starting assumption is that a social tie between two individuals induces them to behave according to some aggregation of their individual preferences. In “[Equilibrium predictions under models of social](#)

preferences,” we consider two leading theories of social preferences (inequity aversion and fairness) that are easily comparable to our approach in the specific interactive strategic situation where the behavioral effects of social ties are evaluated.

More precisely, our current approach is inspired by the existing concept of empathetic preferences as presented by Binmore in (2005): an agent’s empathetic preferences consist in combining his actual own preferences with his preferences when imagining himself in the other agent’s position. In other words, an empathetic agent does not take into account the other’s actual decision, he instead only reasons about his own decision in the other’s position (taking into account the other’s preferences). This concept also refers to the existence of a “veil of ignorance,” as introduced by Rawls in (1971), behind which agents make their decision without knowing in which player’s position they will actually act. As a consequence, such an empathetic behavior can be reduced to simply choosing the corresponding action from the strategy profile that maximizes the group utility.

One should note that our model is also related with Alger and Weibull’s model of Homo Moralis (2012). The difference is that their model requires the game to be symmetric,² whereas our model does not have any restriction and can apply to all sorts of games.

Formally, let us consider two players i and j (note that this model could easily be generalized to an arbitrary number of players). Moreover, let S_i and S_j respectively denote the set of i ’s strategies and the set of j ’s strategies, and $\pi_i(s_i, s_j)$ the payoff function for player i when both i and j respectively play their strategy s_i and s_j . For every $s_i \in S_i$ and $s_j \in S_j$, the *Social Ties* utility function of player i is then given by:

$$U_i^{ST}(s_i, s_j) = (1 - k_{ij}) \cdot \pi_i(s_i, s_j) + k_{ij} \cdot \max_{s'_j \in S_j} U(s_i, s'_j)$$

where $k_{ij} \in [0, 1]$.

The function $U(s_i, s_j)$ stands for the group utility function, which may be characterized by one of the following two well-known principles.

Let us first define a group utility function $U_m(s_i, s_j)$ that satisfies Rawls’ *maximin* criterion (1971).

$$U_m(s_i, s_j) = \min\{\pi_i(s_i, s_j), \pi_j(s_i, s_j)\}$$

This criterion corresponds to giving infinitely greater weight to the benefits of the worse-off person.

As an alternative, one may also consider a function of social welfare $U_s(s_i, s_j)$ that satisfies classical utilitarianism (i.e., by maximizing the total combined payoff of all players).

$$U_s(s_i, s_j) = \pi_i(s_i, s_j) + \pi_j(s_i, s_j)$$

Parameter k_{ij} in the *Social Ties* utility function measures agent i ’s subjective social tie towards agent j . Setting k_{ij} to 0 corresponds to a nonexisting tie (e.g., j is a perfect stranger to i) whereas setting k_{ij} to 1 means that i feels socially very close to j (e.g., j is

² A game is symmetric when all players can switch roles without changing their strategies and the associated payoff.

i 's best friend). In the latter case, one should note that, in the presence of a strong tie with agent j , agent i does not face a strategic problem anymore: indeed, j 's strategy s_j becomes irrelevant to the calculation of i 's utility. Thus, agent i only needs to solve a classical problem of individual decision making by selecting the action from the strategy profile which maximizes the group utility. As a result, i 's action may then be interpreted as "doing the right thing for the group assuming that all other players also do the right thing for the group."

In order to make our approach operational, through the next section we propose a concrete scenario involving strategic interaction among players. The specific game that we consider leads to theoretical predictions under our social ties model that differ both from those obtained under the traditional assumption of self-interested players and by those provided by well-known theories of social preferences. Hence, we think that this example is appropriate to show how our novel approach can explain players' behaviors not captured by leading approaches respectively in traditional and behavioral game theory.

A social ties game

Having previously analyzed the main characteristics of social ties, we now propose the following game that appears to be well suited to study their behavioral effects.

Two colleagues, Alice and Bob, have agreed to cook together and eat in Alice's place, though Bob likes cooking less than Alice. Half an hour before the dinner, Alice makes a phone call to Bob so as to express her interest to go out to her favorite Japanese restaurant, whereas Bob instead suggests to go to a nice Italian restaurant that he recently heard about, with the two restaurants located on opposite sides of the city. Alice then makes it clear that she prefers them cooking rather than Italian food. On the other hand, Bob indicates that, even though he does not particularly enjoy cooking, he still prefers it to the Japanese cuisine. Moreover, Bob reminds Alice of his seafood allergy, which makes eating Japanese more costly for him than eating Italian is costly for Alice. Both individuals also commonly realize that, as the restaurants in question are very popular, they are always overcrowded, which makes it impossible for a person alone to get a table there. Finally, Alice notifies Bob of her intention to go eat outside before the phone call suddenly gets interrupted.

In this scenario, the main questions that arise are the following: assuming Alice and Bob barely know each other, will they manage to meet and eat at the same restaurant that evening? Would they behave differently if they were very close friends instead?

For example, the fact itself that Alice would notify Bob her intention to go eat outside may depend on their social tie. And even the sign of this dependence cannot be easily disclosed through intuition alone. Indeed, on the one hand, a close relationship with Bob should lead Alice to take the "risk" of leaving her place so as to possibly meet him in his preferred restaurant, thereby maximizing their group utility function. On the other hand, the social tie might lead Alice to confirm that they will both eat at her place before the unexpected interruption of the phone call: if she knows that Bob, because of their social tie, is "uncertain" between maximizing the

group utility function (going to his preferred restaurant) and behaving in a fully rational way (going to her preferred restaurant), then she should safely reaffirm that they eat at her place, in order to prevent miscoordination.

In order to formally analyze the dilemma provided by this situation, let us define the corresponding abstract representation, as shown in Fig. 1, which we will call the Social Ties (ST) game. This ST game defines a two player game organized in two stages. During the first stage, only Player *a* (i.e., Alice) is active, and she has to choose between *In* or *Out* (i.e., whether to stay at Alice's place or go eating outside). In the latter case, the outcome is more beneficial to Alice who gets 20 (Bob only gets 10 as it is more costly for him to cook than it is for Alice). On the other hand, if Alice chooses to go eat outside (i.e., *In*), both players enter the second stage of the game that corresponds to a basic coordination game. C_a and C_b define respectively Alice and Bob's choice to go to the Japanese restaurant. Similarly, D_a and D_b define Alice and Bob's choice to go to the Italian restaurant. If both coordinate on the (C_a, C_b) solution, then Alice and Bob get 35 and 5, respectively, and if both coordinate on the (D_a, D_b) solution, then they get 15 (Alice) and 35 (Bob). In any other case (i.e., where they choose different restaurants and are therefore unable to get a table), both players win nothing (0).

As we will detail in the next sections, this game provides an ideal environment to discriminate between different theories of how two socially tied people should act.

One may note that our ST game corresponds to a variant of the Battle of the Sexes (BoS) game with outside option (see Cooper et al. 1993). Indeed, as shown in Fig. 2a, the only difference lies on the symmetrical property within the coordination subgame that we voluntarily removed here: unlike in the BoS game, the lowest payoff is different in the two coordination outcomes (eating Japanese is more costly to Bob than eating Italian is for Alice: $5 \neq 15$). The main motivation to introduce this type of asymmetry is to create some incentives for players to favor the group as a whole (in fact, there is no unique best outcome for the group in a BoS-like subgame). However, one may note that the dilemma introduced by the ST subgame does only concern Alice's preferences: if Alice is self-interested, she will aim at reaching the (C_a, C_b) outcome, whereas if she considers the social welfare of the group, she might wish to reach the (D_a, D_b) outcome. On the other hand, Bob's preference orderings in this

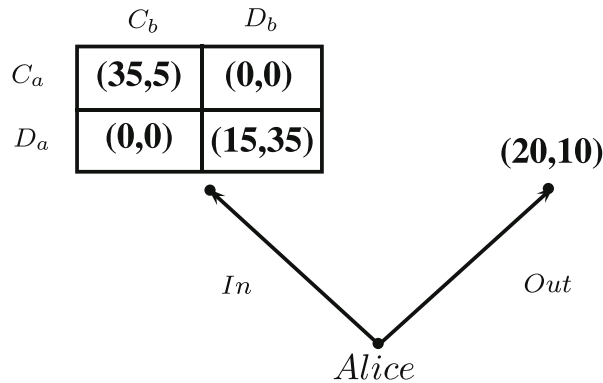


Fig. 1 Social ties game

	C_b	D_b		C_b	D_b		C_b	D_b
C_a	(35,15)	(0,0)	C_a	(35,35)	(0,0)	C_a	(35,15)	(0,0)
D_a	(0,0)	(15,35)	D_a	(0,0)	(15,15)	D_a	(0,0)	(25,25)

a Battle of the Sexes (BoS) game **b** Hi-Lo matching game **c** Dalek subgame

Fig. 2 Existing coordination games

subgame are the same no matter whether he wishes to maximize his self-interest or the social welfare.

Similarly, considering a Hi-Lo-like subgame would not match our current study as it does not offer any dilemma between satisfying self-interest and maximizing the social welfare: in fact, in the Hi-Lo matching game (see Fig. 2b), both players always obtain the same payoff no matter the outcome, which explains the high rate of coordination on the most profitable outcome for both players, independently of whether there exists a tie between them.

One may also notice the similarity of our ST game with the Dalek game presented in (Binmore and Samuelson 1999) (a corresponding subgame is depicted in Fig. 2c). The main difference is that one solution of the coordination subgame ensures perfect equity in the Dalek game. Indeed, as in our case, the Dalek game also introduces some dilemma between maximizing one's self-interest and playing the fairest outcome. However, unlike in our ST game, it does not introduce any dilemma between satisfying self-interest and maximizing the social welfare (i.e., the combined payoffs of every player). Although this game would be interesting to investigate, focusing on it may also make it more difficult to observe the actual effects of social ties on behavior: as a consequence, the absence of any clear incentive to play the fairest solution in the Dalek game may eventually lead to a higher rate of miscoordination, independently of the presence of such ties. On the other hand, the signal of perfect equity in the Dalek game may also appear so strong that it could reinforce the stability of coordinating on the corresponding solution (i.e., (D_a, D_b)), even when no ties are involved.

Equilibrium predictions with self-interested players

To begin with, we analyze the above ST game by assuming that agents are solely own-payoff maximizers.

Through the rest of the paper, we note C_i and D_i to respectively stand for sub-strategies “ C_i if In” and “ D_i if In” in the context of the ST game.

Let us first state that this ST game contains three Nash equilibria (NE) in pure strategies, which are the following:

$$(\text{In}, C_a; C_b), (\text{Out}, C_a; D_b), (\text{Out}, D_b; D_b)$$

Moreover, the ST game also has a Nash equilibrium in mixed strategy, which consists of Alice always playing Out (i.e., with probability 1) and Bob playing D_b with a probability of 3/7. This solution however has to be distinguished from another Nash equilibrium in behavioral strategy, which consists of Alice always playing Out

first (i.e., with a probability of 1) and playing D_a with a probability of $1/8$ in the subgame while Bob then plays D_b with a probability of $7/10$ (Note that this corresponds to the Nash equilibrium in mixed strategy in the subgame). Thus, the respective expected payoffs in both of these cases are 20 for Alice and 10 for Bob. However, it is worth indicating that all NE in mixed or behavioral strategies are simply irrelevant to the ST game: if Alice is willing to randomize in the subgame or believes that Bob will, then she is always better off by playing Out in the first place.

Considering subgame perfect NE, which can be computed through the backward induction method, allows to rule out the solution (Out, C_a ; D_b) even though it is a Nash equilibrium. Indeed, although the prediction to play Out is perfectly rational for Alice, it here relies on the fact that she would not be rational had she played In the first place: given that Bob plays D_b in the subgame, Alice's only rational move would be to play D_a instead of C_a (which corresponds to a Nash equilibrium in the subgame). Note that the previous Nash equilibrium in mixed strategy is ruled out by this principle while the Nash solution in behavioral strategy still remains.

Furthermore, considering the forward induction (FI) principle allows to restrict the previous set of subgame perfect NE to those solutions, which resist the iteration of weak dominance. In the context of our ST game, this leads to the following solution: first Alice's strategy (In, D_a) is weakly (and strictly) dominated by any strategy involving Out. Then Bob's strategy D_b becomes weakly dominated by C_b . Thus Alice's strategies (Out, C_a) and (Out, D_a) are both weakly (and strictly) dominated by (In, C_a ; C_b). Therefore, the unique FI solution, which resists iterated weak dominance, is as follows:

$$(\text{In}, C_a; C_b)$$

Indeed it turns out that fully rational players should play this solution, which can be interpreted as follows: while playing In, Alice signals Bob that she intends to play C_a (if she intended to play D_b , she would have played Out in the first place). Therefore Bob's unique rational move is to play C_b . However, while this interpretation justifies the existence of the above solution, it does not explain why the other backward induction solution is not rational. To continue the argument, let us then consider the solution (Out, D_a ; D_b), which can be interpreted as follows: Alice plays Out because she expects Bob to play D_b in case she had played In. This chain of reasoning is clearly erroneous because Alice's conditional expectation does not match what she would really expect had she *actually* chosen to perform In. Indeed, as shown before, if Alice performs In, Bob's unique rational move is to play C_b , thus no matter what Alice does during the first stage, she cannot expect anything else than Bob playing C_b . Consequently, her unique rational move is to play (In, C_a), and Bob's best response is to play C_b . Moreover, note that, for the same reason, the previous Nash equilibrium in behavioral strategy does not resist this FI argument.

The interesting characteristic that this analysis brings about is that the validity of this FI argument is independent of Bob's preferences. This therefore suggests that such a game introduces some "first-mover" advantage, assuming

that it is common knowledge among them that they both are self-interested agents.

Many studies in the experimental economic literature have provided support to this FI argument (see, e.g., Balkenborg 1994; Brandts and Holt 1989, 1995; Cachon and Camerer 1996; Cooper et al. 1992, 1993; Shahriar 2009; Van Huyck John et al. 1993).

One of the first papers in this direction is Brandts and Holt (1989). Cooper et al. (1992) investigate a coordination game with two Pareto-ranked equilibria and report that a payoff-relevant outside option changes play in the direction predicted by FI. Van Huyck John et al. (1993) report the success of FI in a setup in which the right to participate in a coordination game is auctioned off prior to play. Cachon and Camerer (1996) investigate a setup in which subjects may pay a fee to participate in a coordination game with Pareto-ranked equilibria. They report that play is consistent with FI.

While many experiments support the fact that people's strategic behavior relies on the FI argument (see, e.g., Cooper et al. 1992, 1993; Van Huyck John et al. 1993), there is also contrary evidence. In (1993), Cooper et al. obtain the FI solution when it coincides with a dominance argument but the same outcome is predicted when FI makes no prediction. Brandts and Holt (1995) also show that the FI is a good prediction only if it coincides with a simple dominance argument. In (2003), Brandts et al. find evidence against FI in an industrial organization game.

Other work have shown that the temporal factor of the game is relevant to FI reasoning. In Cooper et al. (1993) and Huck and Muller (2005), the FI solution predicts well subjects' behavior in an experimental game in extensive form, but does poorly when subjects are presented with the normal form game. A similar problem seems to arise in Caminati et al. (2006) who analyze games similar to ours but who work essentially with the normal form.

However, all these works consider games that are slightly different from the interactive strategic situation on which we focus in this paper. One may then wonder whether the asymmetry introduced in our ST game does alter the game theoretical prediction.

Equilibrium predictions under models of social preferences

In this section, we reinterpret our ST game through the use of well-known economic theories of social preferences and analyze players' equilibrium behavior under these theories. In fact, these models allow one to consider not only the self-interested motivations of the players, but also their social motivations, which may then be particularly important in the context of social ties. In other words, a player's utility is not characterized by his own material payoff only, but also those of the other players. We choose to focus on the concepts of inequity aversion and fairness, which seem to be the most relevant to our ST game. Other models of intentions-based fairness and reciprocity (see, e.g., Rabin 1993b) do not appear to be suitable to such a coordination game. Apart from the problem of multiple equilibria in beliefs that characterizes such belief-

dependent approaches, it would be difficult to unambiguously define what is “kind” and what is “unkind” in the players’ strategy set, by using only first and second order beliefs.

Theory of inequity aversion

In the models proposed by Fehr and Schmidt (1999) and Bolton and Ockenfels (2000), players are assumed to be intrinsically motivated to distribute payoffs in an equitable way: a player dislikes being either better off or worse off than another player. In other terms, utilities are calculated in such a way that equitable allocations of payoffs are preferred.

Formally, consider two players i and j and let $x = \{x_i, x_j\}$ denote the vector of monetary payoffs. According to Fehr and Schmidt’s *inequity aversion* model, the utility function of player i is given by:

$$U_i^{IA}(x) = x_i - \alpha_i \cdot \max\{x_j - x_i, 0\} - \beta_i \cdot \max\{x_i - x_j, 0\}$$

where it is assumed that $i \neq j$, $\beta_i \leq \alpha_i$ and $0 \leq \beta_i < 1$.

The two parameters can be interpreted as follows: α_i parametrizes the distaste of person i for disadvantageous inequality while β_i parametrizes the distaste of person i for advantageous inequality. One should note that setting these parameters to zero defines some purely self-interested agent. The constraints imposed on the parameters are meant to ensure that players cannot distaste advantageous inequality more than disadvantageous inequality in order to be realistic.

Clearly, applying such a model to our current ST game can literally transform its whole structure, depending on the values assigned to parameters α_i and β_i . Let us then perform a game theoretical analysis that involves such inequity aversion parameters.

The main observation that can be made is about the effects of Alice’s preference ordering on her behavior. In fact, assuming that $\beta_a \leq \alpha_a$, then Alice will never play the strategy (In, \mathbf{D}_a), no matter how inequity averse she is:

- if $\beta_a < 3/4$ and $\alpha_b < 1/6$, then Alice and Bob’s optimal behavior remains as if they were self-interested (i.e., the FI argument still holds). Thus, Alice’s unique rational strategy is to play (In, \mathbf{C}_a) while Bob will rationally play \mathbf{C}_b .
- if $\beta_a < 3/4$ and $\alpha_b \geq 1/6$, then Alice is always better off by playing (Out, \cdot): the coordination subgame yields a unique Nash equilibrium (i.e., (D_a, D_b)), which is strictly dominated by strategy (Out, \cdot).
- if $\beta_a \geq 3/4$, then Alice is always better off by playing (Out, \cdot): for any $\alpha_a \geq \beta_a$, any outcome from the coordination subgame is strictly dominated by playing (Out, \cdot) (see Fig. 3 for an example).

The main result of this analysis is that the value of α_a and β_b are irrelevant to defining Alice and Bob’s optimal behavior. In other words, only Alice’s distaste about advantageous inequality can affect her preference ordering in the ST game. Similarly, only Bob’s distaste about disadvantageous

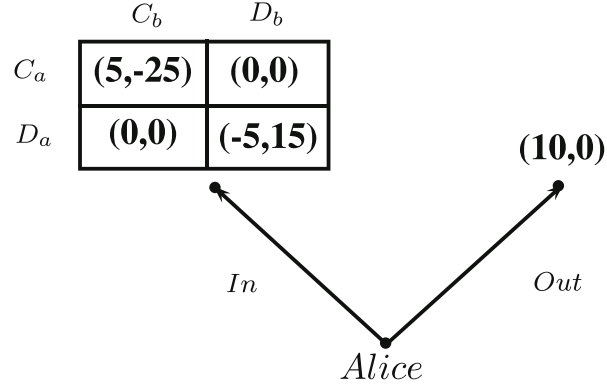


Fig. 3 Transformed ST game with extremely inequity averse players ($\alpha_a = \beta_a = \alpha_b = \beta_b = 1$)

inequality can affect his decision. One should also note that inequity aversion does not keep the “first mover” advantage mentioned in the previous section: Alice’s first move does signal Bob not only about her low level of inequity aversion, but also about her expectation of Bob’s low level of inequity aversion. That means that if she plays In, then the resulting outcome is entirely depending on Bob’s level of inequity aversion (either (In, C_a ; C_b) or (In, C_a ; D_b) will be played).

The sets of NE and of Subgame Perfect Nash Equilibria (SPE), in the context of the ST game played with inequity aversion, are summarized in Table 1 (note that FI is irrelevant in this case because the SPE is unique for every vector (α_i, β_i) , with $i \in \{a, b\}$ and $\alpha_i, \beta_i \in [0, 1]$).

Theory of fairness

Let us now consider another type of social preferences model that in turn relies on the notion of fairness. In (2002), Charness and Rabin propose a specific form of social preference they call *quasi-maximin* preferences. In their model, group payoff is computed by means of a social welfare function which is a *weighted* combination of Rawls’ *maximin* and of the utilitarian welfare function (i.e., summation of individual payoffs) (see, Charness and Rabin 2002, p. 851).

Table 1 Equilibrium solution concepts for inequity averse agent(s) ($\beta_a \geq 3/4$ or $\alpha_b \geq 1/6$)

NE	SPE
(Out, C_a , C_b)	(Out, C_a , C_b) if $\alpha_b < 1/6$
(Out, D_a , D_b)	(Out, D_a , D_b) if $\beta_a < 3/4$
(Out, C_a , D_b)	(Out, C_a , D_b) if $\alpha_b \geq 1/6$ and $\beta_a \geq 3/4$
(Out, D_a , C_b)	

Formally, consider two players i and j and let $x = \{x_i, x_j\}$ denote the vector of monetary payoffs. According to Charness and Rabin's *fairness* model, the utility function of player i is given by:

$$U_i^F(x) = (1-\lambda) \cdot x_i + \lambda \cdot [\delta \cdot \min\{x_i, x_j\} + (1-\delta) \cdot (x_i + x_j)]$$

where $\delta, \lambda \in [0, 1]$. Moreover, the two parameters can be interpreted as follows: δ measures the degree of concern for helping the worst-off person versus maximizing the total social surplus. Setting $\delta=1$ corresponds to a pure “*maximin*” (or “Rawlsian” criterion), while setting $\delta=0$ corresponds to total-surplus maximization. Parameter λ measures how much player i cares about pursuing the social welfare versus his own self-interest. Setting $\lambda=1$ corresponds to purely “disinterested” preferences, with player i caring no more (or less) about her own payoffs than others’, while setting $\lambda=0$ corresponds to pure self-interest.

As for the previous model, the parameters δ and λ can considerably change the structure of the ST game, which is why we propose a new game theoretical analysis involving such fair agents.

The first observation is that while fairness may slightly alter Bob’s preferences, the (In, \mathbf{D}_a ; \mathbf{D}_b) outcome always remains the best option: the only difference with the classical model is that he may come to prefer the (In, \mathbf{C}_a ; \mathbf{C}_b) outcome to the (Out, \cdot) outcome when $\delta < 2/3$ and $\lambda > 1/3$.

Similarly, Alice’s preferences also get affected by such notion of fairness. The main result is that a new FI solution may emerge through such a social preferences model. In particular:

- If $\lambda < 1/2$, then Alice may still play the FI solution strategy as predicted by traditional game theory (i.e., (In, \mathbf{C}_a)), depending on the value of δ .
- If $1/2 \leq \lambda \leq 3/4$, then no prediction can be made without considering probabilistic beliefs: both NE in pure strategies in the subgame are always at least as good for Alice as playing (Out, \cdot).
- If $\lambda > 3/4$ and $\delta > 2/3$, then Alice may play a FI solution strategy (i.e., (In, \mathbf{D}_a)) that mainly relies on her other regarding preferences (see Table 2): solution (In, \mathbf{D}_a ; \mathbf{D}_b) indeed becomes preferred to playing (Out, \cdot), which is preferred to solution (In, \mathbf{C}_a ; \mathbf{C}_b) (see Fig. 4 for an example).

Moreover, one should note that, as for the original version of the game (see “Equilibrium predictions with self-interested players”), the Out option for Alice always dominates the Nash equilibrium in mixed strategies from the coordination subgame, no matter what the values of λ and δ are.

Table 2 Equilibrium solution concepts for fair agents ($\lambda >> 3/4$ and $\delta >> 2/3$)

NE	SPE	FI
(In, \mathbf{D}_a , \mathbf{D}_b)	(In, \mathbf{D}_a , \mathbf{D}_b)	(In, \mathbf{D}_a , \mathbf{D}_b)
(Out, \mathbf{C}_a , \mathbf{C}_b)	(Out, \mathbf{C}_a , \mathbf{C}_b)	
(Out, \mathbf{D}_a , \mathbf{C}_b)		

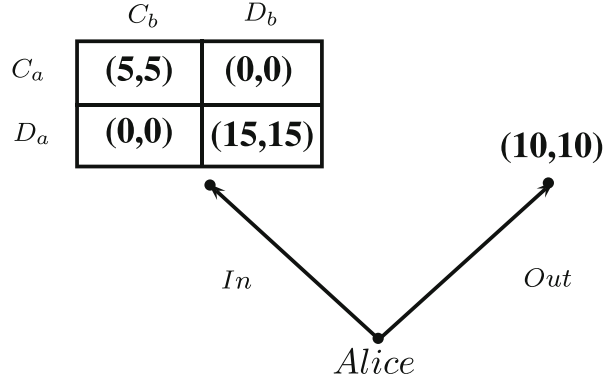


Fig. 4 Transformed ST game for extremely fair agents ($\lambda=\delta=1$)

The above analysis suggests that the ST game may in fact contain two distinct focal points for the players, which can be identified by the two possible FI solutions. Therefore, one can state that the current ST game yields a unique social-welfare equilibrium³ if and only if players have either some strong self-interested preferences ($\lambda \ll 1/5$) or some strong other-regarding preferences ($\lambda \gg 3/4$ and $\delta \gg 2/3$). In the latter case, one should note that the players' sensibility to the *maximin* principle needs to “dominate” that of the utilitarian welfare function.

The sets of NE, of SPE, and of FI solutions, in the context of the ST game played by fair agents, are shown in Table 2.

Equilibrium predictions under our model of social ties

Similarly to the theories of social preferences considered in the previous section, our main claim is that the strength of the social tie existing between two players has some important effects on their preferences (and consequently on their expected behavior), as suggested in “[A theory of how to model social ties](#).” However, as we believe that the type of *payoff transformation* used in our model is more appropriate to the context of social ties, we illustrate through this section how it can disagree with the previous concepts from “[Equilibrium predictions with self-interested players](#)” and “[Equilibrium predictions under models of social preferences](#)” in terms of equilibrium predictions. More specifically, let us apply our model of social ties to the above ST game.

Given the subjective social ties parameters k_{ab} and k_{ba} , and the constant values $k_1, k_2, k_3 \in [0,1]$ such that $k_1 \leq k_2 \leq k_3$, we have:

- if $k_{ab} > k_3$, then the unique rational play for Alice and Bob is to coordinate on (In, $\mathbf{D}_a; \mathbf{D}_b$), independently of k_{ba} .

³ The social welfare equilibrium introduced by Charness and Rabin (2002, p. 852) corresponds to a Nash equilibrium for some given values of δ and λ .

	C_b	D_b
(In, C_a)	(5,10)	(5,15)
(In, D_a)	(15,10)	(15,15)
(Out, C_a)	(10,10)	(10,15)
(Out, D_a)	(10,10)	(10,15)

Fig. 5 Transformed ST game for socially tied agents ($U=U_m$ and $k_{ab}=k_{ba}=1$)

- if $k_{ab} > k_1$ and $k_{ba} > k_2$, then, as in the previous case, the unique Nash equilibrium is again for both players to coordinate on the $(In, D_a; D_b)$ outcome (see Fig. 5 for an example).
- if $k_{ab} \leq k_1$ and $k_{ba} > k_2$, then Alice should play (Out, \cdot) , in response to Bob playing D_b in the subgame.
- if $k_{ab} \leq k_1$ and $k_{ba} \leq k_2$, then Alice and Bob should follow FI reasoning and play $(In, C_a; C_b)$. In this case, as the strategic structure of the game remains as in its original version, the game-theoretic analysis from “[Equilibrium predictions with self-interested players](#)” still applies.
- if $k_1 < k_{ab} \leq k_3$ and $k_{ba} \leq k_2$, then both players are unable to coordinate on a unique Nash equilibrium outcome in the subgame: both (C_a, C_b) and (D_a, D_b) are NE. As a result of such indecision in the subgame, Alice’s optimal strategy is (Out, \cdot) .

In the above analysis, the constant values for k_1 , k_2 , and k_3 depend on whether the players follow the *maximin* function (i.e., $U=U_m$) or the utilitarianism principle (i.e., $U=U_s$). Table 3 provides the corresponding constant values for each of these types. Moreover, this analysis considers the most general case where ties may be unilateral (i.e., where $k_{ab} \neq k_{ba}$): our current model therefore allows to state that, for instance, Alice feels close to Bob while Bob does not feel close to Alice.

One should note from this interpretation that Alice will always play her strategy (In, D_a) whenever $k_{ab} > k_3$, and similarly, Bob will always play his strategy D_b

Table 3 Constant tie values for each type of player

Constant (k_i)	Egalitarianism ($U=U_m$)	Utilitarianism ($U=U_s$)
k_1	1/2	1/5
k_2	1/2	1/3
k_3	4/5	7/9

whenever $k_{ba} > k_2$. Such an observation, combined with the fact that $k_3 > k_2$, indicates that Alice's decision is more restrictive than Bob's: if $k_2 < k_{ab}$, $k_{ba} < k_3$, then Alice needs to take Bob's decision into account in order to make her decision, whereas Bob will play \mathbf{D}_b independently of Alice's action. As a consequence, introducing social ties in the context of this game may allow Bob's threat of playing \mathbf{D}_b to become more credible to Alice's eye. In other words, social ties may simply turn Alice's first mover advantage (as suggested in Huck and Muller 2005) in Bob's favor.

One should also note the distinction between using utilitarianism or egalitarianism in the tie utility function. As shown in Table 3, utilitarianism allows to coordinate more easily on the (In, \mathbf{D}_a ; \mathbf{D}_b) outcome than egalitarianism (i.e., k_1 , k_2 , and k_3 have lower values when $U = U_s$). On the other hand, players following egalitarianism are expected to coordinate in the subgame more often than in the case of utilitarianism (i.e., when $U = U_m$, Alice should play (Out, \cdot) if and only if $k_{ab} = k_1 = k_2 = 1/2$).

However, one may state that the formulation of social ties from “A theory of how to model social ties” is too general with respect to the concept presented in “A definition of social ties.” In fact, according to Statement 2.01, we assume that a social tie is restricted to be bilateral, which may not always be the case in the above analysis. In order to match this criterion, one then needs to add the following constraint:

$$k_{ab} = k_{ba}$$

In this case, assuming $k = k_{ab} = k_{ba}$, coordination on the (In, \mathbf{D}_a ; \mathbf{D}_b) outcome is reached only when $k > k_2$. Similarly, coordination on the (In, \mathbf{C}_a ; \mathbf{C}_b) outcome is reached only when $k < k_1$. On the other hand, Alice will play (Out, \cdot) whenever $k_1 \leq k \leq k_2$, as miscoordination would be expected in the subgame. As this constraint simplifies decision making for both players in the ST game, one should note that it also removes any opportunity for Bob to exploit Alice: whenever $k_2 < k_{ab}$, Alice will always play (In, \mathbf{D}_a) independently of Bob's decision.

More specifically, about determining Alice and Bob's dinner plan, this analysis leads to the following interpretation. On the one hand, if the existing social relationship is sufficiently weak between Alice and Bob, then Bob should be influenced by Alice's intention of going to her favorite Japanese restaurant and, consequently, he should choose \mathbf{C}_b in order to maximize her own payoff. On the other hand, in the presence of a sufficiently strong social tie between them, Alice should be influenced by Bob's intention of doing what is best for the group and, consequently, she should choose \mathbf{D}_a in order to maximize the group payoff.

Furthermore, in the particular case of an intermediate measure of social ties between Alice and Bob, each individual may then become uncertain about the other's choice of either acting fully rationally and go to Alice's favorite restaurant, or doing what is best for the group and go to Bob's favorite restaurant. In response to such a high risk of meeting at the wrong restaurant, Alice should then reaffirm her intention to stay eat at her place with Bob.

However, one should note that the above constraint on both players' subjective social ties (i.e., $k_{ab} = k_{ba}$) suggests the existence of a common scale for measuring such ties. In fact, in order for Alice to determine the actual social tie level between Bob and herself, then she must first make sure that her notions of the weakest and highest possible ties are the same as Bob's (e.g., does “being best friends” have the same meaning for both of them?). For example, it is fair to say that everybody does not share the same level of tie with a complete stranger (some are genuinely more

cooperative than others). One may then argue that the non-easy task of normalizing one's social ties scale is already part of what defines the social tie itself: indeed the higher the social tie value, the more the concerned individuals are likely to share the same social ties scale (e.g., if I feel that we are best friends, but I ignore what "being best friends" means to you, then I will not risk being exploited by you, and my social tie with you will consequently remain weak). As a result, this analysis follows our definition of social ties from "[A definition of social ties](#)" as it suggests that social ties are intrinsic psychological factors that are influenced by the agents' epistemic states (the more I know about you, the more my measure of our social tie is reliable).

Players as team-directed reasoners

Our proposed model of social ties from "[A theory of how to model social ties](#)" appears to share some common properties with another well-known concept that also relies on group identification: team reasoning. We therefore provide, in this section, a detailed analysis of the ST game through the various theories of team reasoning as a means to illustrate the common characteristics as well as the differences that exist with our model of social ties.

One should note that, in the context of the ST game, considering collective utility functions (see, e.g., the classical utilitarianism and the *maximin* principle from "[A theory of how to model social ties](#)") from the players' individual viewpoint can lead to a transformed game similar to that depicted in Fig. 4 from "[Theory of fairness](#)". In this case, the resulting subgame in Fig. 4 has similar properties as the well known Hi-Lo matching game: as both players have the same preferences over outcomes, they indeed benefit if and only if they coordinate with each other in the subgame. However, their subsequent payoffs depend on which action they do coordinate on. The interesting property of this transformed subgame is that it introduces a dilemma that even economic theory cannot solve. However, while game theory is indeed unable to predict any particular outcome (i.e., both coordinated outcomes of the subgame are Nash solutions), it is shown in (Bacharach 2006) that people would tend to coordinate on the action that leads to the most rewarding outcome for both, i.e., (D_a, D_b) .

In order to interpret such intuitive behavior, some theorists have proposed to incorporate new modes of reasoning into game theory. For instance, starting from the work of Gilbert (1989) and Regan (1980), some economists and logicians (Lorini 2011) have studied team reasoning as an alternative to the best-response reasoning assumed in traditional game theory (Bacharach 1999; Colman et al. 2008; Sugden 2000, 2003). Team-directed reasoning is the kind of reasoning that people use when they perceive themselves as acting as members of a group or team (Sugden 2000). That is, when an agent i engages in team reasoning, he identifies himself as a member of a group of agents S and conceives S as a unit of agency acting as a single entity in pursuit of some collective objective. A team reasoning player acts for the interest of his group by identifying a strategy profile that maximizes the collective payoff of the group, and then, if the maximizing strategy profile is unique, by choosing the action that forms a component of this strategy profile.

Furthermore, as suggested in (Hakli et al. 2010), the concept of team reasoning also refers to Tuomela’s *I-mode/we-mode* distinction from (Tuomela 2010). According to Tuomela, the *I-mode* consists of reasoning as a private person according to two possible principles: an agent reasoning in *plain I-mode* will seek to satisfy self-interest as suggested by classical economic theory (this corresponds to the type of reasoning underlying the analysis from “Equilibrium predictions with self-interested players”). On the other hand, an agent reasoning in *pro-group I-mode* is concerned with promoting the group’s interests, and as a result will make a decision with the individual intention to maximize the group utility. As an example of such a benefactor behavior (as called by Bacharach in 1999), one may consider the existing theories of social preferences presented in “Equilibrium predictions under models of social preferences” (e.g., the concept of social welfare equilibrium presented in (Charness and Rabin 2002) illustrates this type of thinking). It is clear however that even such *pro-group I-mode* thinking, which relies on preference transformation, fails to predict some very intuitive behavior such as in the well known Hi-Lo matching game, as shown in (Bacharach 2006; Colman et al. 2008) (see Fig. 4 for a similar example). The alternative concept of *we-mode* reasoning instead relies on what Bacharach calls agency transformation, which consists of conceiving the situation not as a decision making problem for individual agents (cf. the *I-mode*), but as a decision making problem for the group conceived as an agent.

Let us now perform a detailed analysis of team reasoning applied to the ST game according to both Sugden and Bacharach’s different theories.⁴ Figure 6 illustrates a representation of the ST game from the group’s viewpoint when considering the *maximin* principle as the group utility function. In this case, the transformed ST game considers a unique player, which corresponds to the group $\{Alice, Bob\}$. We indicate with $s_a s_b$ any group’s strategy. One therefore notes that the best strategy for the group is to always play $(In, D_a; D_b)$.

First, according to Sugden’s theory (Gold and Sugden 2007; Sugden 2003), a simple epistemic interpretation of a team reasoning (from Alice’s viewpoint) in the current ST game can be the following (for $U \in \{U_s, U_m\}$):

Statement 8.01: *If Alice believes that:*

- *She is a member of the group $\{Alice, Bob\}$.*
- *It is common knowledge among Alice and Bob that both identify with $\{Alice, Bob\}$.*
- *It is common knowledge among Alice and Bob that both want the value of U to be maximized.*
- *It is common knowledge among Alice and Bob that $(In, D_a; D_b)$ uniquely maximizes U . Then she should choose her strategy (In, D_a) .*

Following Statement 8.01, it is then clear that if Alice shares her beliefs with Bob, then the resulting outcome will be $(In, D_a; D_b)$; i.e., Bob will similarly choose the corresponding option D_b). According to Sugden (2003), a player has reason to act as a team member and to choose the action that forms a component of the strategy profile

⁴ A more detailed general comparison of Sugden and Bacharach’s theories of team reasoning can be found in (Gold 2012).

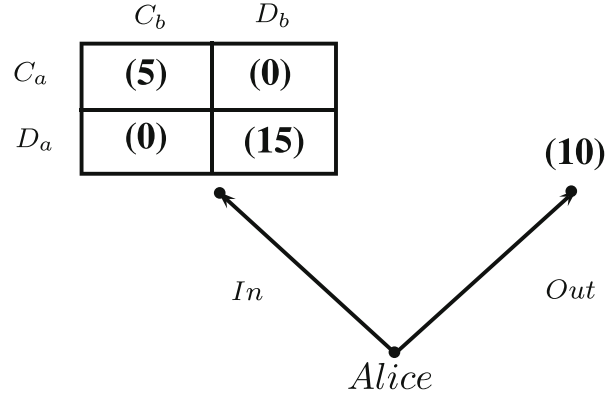


Fig. 6 ST game from the group's viewpoint (with $U=U_m$)

maximizing collective payoff, conditional on assurance that the other players also act as team members. That is, to act as a member of a team, one must be confident that the other players act as members too. More fundamentally, “[...] team reasoning does not generate reasons for choice unless each member of a team has reason to believe that there is *common reason* to believe that each member of the team endorses and acts on team reasoning [...]. This is a condition of assurance” (Tajfel 1970, pp. 176–177). In other words, the main characteristics of Sugden’s theory is that team reasoning relies on strong epistemic foundations and is very restrictive in that matter: an agent will not take the risk to team reason and be “suckered” by other agents who do not.

Let us now consider Bacharach’s theory of team reasoning as an alternative to Sugden’s previous interpretation, which, as shown in (Hakli et al. 2010), yields the same action recommendations as Tuomela’s *we-mode* reasoning in any game-theoretic situation. In (1999), Bacharach introduces the concept of *unreliable team interaction*, which corresponds to a game structure in which there is a probability that a given player identifies with a team and chooses the action which maximizes the team benefit (i.e., the player plays in the *we-mode*), and another probability that the player is a self-interested agent who tries to maximize his own benefit (i.e., the player plays in the *I-mode*). In this sense, the interaction is “unreliable” because there is no certainty that a player will reason and act as a team member. A team member will then act according to his expected utility based on what others will do (including players who do not team reason). In such an *unreliable team interaction*, Bacharach also introduces the notion of a team protocol, which consists of specifying a strategy for every player when identifying with each team (or in each mode). As an example, the protocol $(C_a, C_b, \text{ and } D_a D_b)$ specifies that Alice and Bob will respectively play C_a and C_b if in *I-mode*, and will respectively play D_a and D_b if in *we-mode*. Through this concept of team protocol, Bacharach differentiates an agent’s behavior depending on whether he identifies with the group or not. The players then reach a protocol equilibrium if and only if, given the probability ω that each player reasons in *we-mode*, neither players nor the group can increase its expected utility by individually deviating from it. In other words, such a protocol equilibrium may simply be interpreted as a Nash equilibrium in the extended game where the group {Alice,

Table 4 Protocol equilibria (with $X_a \in \{\text{Out}, \text{In } \mathbf{C}_a, \text{In } \mathbf{D}_a\}$ and $X_b \in \{\mathbf{C}_b, \mathbf{D}_b\}$)

Probability <i>we-mode</i>	Protocol equilibria (for $U \in \{U_m, U_s\}$)
$\omega=0$	(In $\mathbf{C}_a, \mathbf{C}_b, X_a X_b$) (Out, $\mathbf{D}_b, X_a X_b$)
$0 < \omega \leq 3/4$	(In $\mathbf{C}_a, \mathbf{C}_b, \text{In } \mathbf{C}_a \mathbf{C}_b$) if $U = U_s$ (In $\mathbf{C}_a, \mathbf{C}_b, \text{Out } \mathbf{C}_b$) if $U = U_m$ (Out, $\mathbf{D}_b, \text{In } \mathbf{D}_a \mathbf{D}_b$)
$3/4 < \omega \leq 8/9$	(In $\mathbf{C}_a, \mathbf{C}_b, \text{In } \mathbf{C}_a \mathbf{C}_b$) if $U = U_s$ (Out, $\mathbf{D}_b, \text{In } \mathbf{D}_a \mathbf{D}_b$)
$8/9 < \omega < 1$	(Out, $\mathbf{D}_b, \text{In } \mathbf{D}_a \mathbf{D}_b$)
$\omega=1$	($X_a, X_b, \text{In } \mathbf{D}_a \mathbf{D}_b$)

Bob becomes an extra player. Such a property illustrates a major difference between Bacharach’s theory and our model of social ties. In fact, while an *unreliable team interaction* requires to consider additional players in the game in order to perform strategic reasoning, each of which corresponding to a combination of individual players (cf. the concept of a team protocol), our model of ties instead only leads to a modification of each individual’s utility, leaving the game structure in its original version (i.e., the sets of players and strategies remain unchanged).

Table 4 describes the sets of Protocol equilibria for each probability value ω that each player reasons in *we-mode* in the ST game.

One can see from Table 4 that, when the agents reason in *I-mode* (i.e., $w=0$), the set of protocol equilibria matches that of NE. Conversely, if both agents play in *we-mode* (i.e., $w=1$), then the only equilibrium is to play the solution (In, $\mathbf{D}_a, \mathbf{D}_b$). Moreover, one can note that the *maximin* principle requires a lower probability of *we-mode* reasoning ($w > 3/4$) than the utilitarianism principle ($w > 8/9$) in order to converge to this unique solution. In other words, the utilitarianism principle requires a stronger identification with the same group in order to achieve coordination. One should note that team reasoning agrees with our model of social ties regarding the behavioral predictions in the context of the ST game.

However, the main limitation of Bacharach’s theory is that it does not clarify what the probabilistic distribution ω stands for in the definition of an *unreliable team interaction* structure. In fact, while such probabilities may depend on some intrinsic features of the game such as the payoff structure, they may also reasonably be determined by some pre-existent social relationships between the players: two strongly (respectively weakly) tied individuals may indeed each have a high probability of being in *we-mode* (resp. *I-mode*) in situations like the above ST game.

To further the analysis, let us note that the concept of *unreliable team interaction* can be seen as a special type of incomplete information games⁵ where the only uncertainty one can have is regarding the level to which other players identify with different groups (e.g., agent i may identify with the group $\{i, j\}$ with probability ω or with the group $\{i\}$ with probability $1-\omega$). In other words, this theory relies on the

⁵ Note that we do not refer to the usual Bayesian game as defined by Harsanyi here. It is indeed shown in (Hakli et al. 2010) that a Bayesian game generated from Bacharach’s unreliable team interaction structure does not yield the same action recommendation.

assumption that every agent identifies with a unique team at a given time, which is a strong assumption. This observation therefore raises the issue of the endogenous determination of the mode of reasoning (i.e., *I-mode/we-mode*). In fact, a fundamental point in Bacharach's original theory, in contrast to Sugden's theory, is that the determination of mode of reasoning is a psychological matter, prior to any rational choice, and such a process is based on frames. A frame, as first introduced in Bacharach and Bernasconi (1997) through the Variable Frame Theory, can be defined as a set of concepts that an agent uses when thinking about a decision problem: a person may then start to we/I-reason only if he has "we"/"I" concepts in his frame, which leads him to answer the corresponding question "What shall we/I do?". While Bacharach's theory assumes an agent can only use one frame at once (i.e., an agent cannot reason in *I-mode* and *we-mode* at the same time), it is suggested in Smerilli (2010) that some vacillation between different frames may actually occur in one's mind when facing a decision problem. The corresponding model indeed defines the probability ω as a function of the probability of vacillating from *we-mode* to *I-mode* and the probability of vacillating from *I-mode* to *we-mode*. Applying such a model to our current ST game leads to the following interpretation: if Alice and Bob start by we-reasoning, there will be a unique we-equilibrium (In, \mathbf{D}_a ; \mathbf{D}_b), which is not a Nash equilibrium. So if Alice starts with *we-mode* reasoning, she will not be happy with the result and move away from this equilibrium (e.g., by playing (Out, ·)). If instead Alice starts with the *I-mode*, both individuals shall not be happy (both (In, \mathbf{C}_a ; \mathbf{C}_b) and (Out, \mathbf{D}_a ; \mathbf{D}_b) are dominated by the previous "we" solution (In, \mathbf{D}_a ; \mathbf{D}_b)). In this case there can be a continuous switching or vacillation from one frame to another. As shown in (Smerilli 2010), this interpretation is similar to that of the well known prisoner's dilemma, which therefore suggests that miscoordination should prevail in the ST subgame.⁶ Consequently, one could conjecture that reinforcing social ties between individuals in the ST game decreases the probability p of vacillating between *we-mode* and *I-mode*, while increasing the probability q of vacillating between *I-mode* and *we-mode*.

Why team reasoning cannot express gradual social ties

Following the previous analysis of team reasoning, one may however wonder whether this interpretation, and more generally Sugden and Bacharach's theories, are actually adequate to interpret the effects of social ties, as they clearly forbid the possibility that an agent is reasoning in two different modes at the same time. Indeed, in some unpublished work (1997), Bacharach allows for the existence of some "superordinate" frame where an agent can see the problem from both the 'I' and the 'we' perspective, even though he states that those perspectives cannot hold simultaneously. However, this interpretation, which is assumed in Sugden and Bacharach's theories, does clearly not allow to capture the fact that one may identify with a given group *up to a certain degree*. The need for such a gradual group identification is justified by the various social features that may simultaneously define

⁶ Experimental results in the Prisoner's dilemma have shown that the cooperation rate varies between 30 and 40 % (see, e.g., Shafir and Tversky 1992).

one's social identity, as suggested by our basic definition of social ties from “[A definition of social ties](#).” For example, two individuals may consider political orientations (e.g., being a Democrat) and religion (e.g., being a Catholic) as very important social features. In this case, it is reasonable to state that the social tie between them if they share both of these features (e.g., both are Democrat and Catholic) is *stronger* than if they share only one of those (e.g., both are Democrat, but one is Catholic while the other is Muslim), which is itself assumed to be *stronger* than sharing none of them (e.g., one is Republican and Catholic while the other is Democrat and Muslim).

In order to illustrate more formally the differences existing between our model of social ties from “[A theory of how to model social ties](#)” and the concept of team reasoning from “[Players as team-directed reasoners](#),” let us consider a simple concrete two player game where *Player (i)* can choose between three options: *A*, *B*, and *C*. In such a scenario, each player's payoff is determined uniquely from these options according to Table 5 (for simplicity, *Player (j)* has no control over the outcome).⁷ Note that the pair's payoff function can then follow either utilitarianism (i.e., sum of individual payoffs) or the *maximin* principle (i.e., minimum of individual payoffs).

Applying team reasoning to this particular situation leads to the following predictions: *Player (i)* will play *A* if reasoning in *I-mode* (*Player (i)* is then self-interested), and *Player (i)* will play *B* if reasoning in *we-mode* (*Player (i)* then identifies with the group). As a consequence, according to both Bacharach and Sugden's theories of team reasoning (and independently of whether the group utility function follow utilitarianism or the *maximin* principle), *Player (i)* will never happen to choose *C*. Indeed, while Sugden's theory does clearly not allow for such gradual team identification, Bacharach's theory cannot provide an *unreliable team interaction* structure with a probability ω of identifying with the group ($0 \leq \omega \leq 1$) that specifies this outcome to occur.

On the other hand, considering the same game through our model of social ties from “[A theory of how to model social ties](#)” leads to a different interpretation: in this case, *Player (i)* will select *A* if both players are extremely close to each other (e.g., $k_{ij}=k_{ji}=1$), and *Player (i)* will select *B* if they instead are perfect strangers (e.g., $k_{ij}=k_{ji}=0$). However, if both players are neither best friends nor perfect strangers but, say, simple acquaintances (e.g., $k_{ij}=k_{ji}=0.5$), then *Player (i)* will choose *C* (assuming either utilitarianism or the *maximin* principle as the group utility function), as a compromise between being self-interested and altruistic. One can therefore observe that the concept of team reasoning has limited expressive power in the context of social ties, as all the theories presented “[Players as team-directed reasoners](#)” are unable to make such an intuitive prediction.

In fact, we claim that our model of social ties is more general than team reasoning in the sense that it allows for a gradual measure of group identification, which, we believe, is an important requirement to capture the actual ongoing behavior under the effects of social ties, as suggested by “[A definition of social ties](#)”: according to our model, an agent may indeed partially identify with the group while remaining partly

⁷ The game presented here is overly simplistic as a means to illustrate the above point, which could of course also be found in more classical types of social interactions.

Table 5 Simple dictator game

Player (<i>i</i>)'s option	Payoffs			
	Player (<i>i</i>)	Player (<i>j</i>)	Collective (sum)	Collective (min)
<i>A</i>	8	0	8	0
<i>B</i>	5	7	12	5
<i>C</i>	7	4	11	4

self-interested. More generally, our model allows to represent the fact that a person is socially tied with another individual *up to a certain degree*.

Furthermore, the characteristics of our model of social ties seem even more suitable to realistically represent more flexible and heterogenous multi-player interactions where different coalitions might be formed (depending on the way the agents are socially tied with one another). Such complex games indeed justify the need for a continuous measure of group identification: as social ties are defined as independent values from an individual to another, an agent may then be similarly tied with various individuals at the same time. Consider, for example, a scenario where one faces the dilemma between cooperating with a close friend, and cooperating with a family member. In this case, assuming the friend and the family member do not know each other, it does clearly not make sense for one to reason as a unique team including all individuals. As a result of identifying with either of the subgroups (i.e., either the one including the friend or the one including the family member), the theories of team reasoning would then predict that one will choose over these two options, even though a more egalitarian solution might exist that would satisfy more equally all players. Under the same condition, our model of social ties would instead suggest to select the latter solution. We however postpone the detailed study of such complex situations to future work.

Further hypotheses

As previously mentioned, the main goal of our ST game is to investigate whether social ties affect social preferences. According to our model of social ties presented in “A theory of how to model social ties,” our first hypothesis is that social ties do not correlate with inequity aversion. Indeed, while our model of ties supports coordination on the $(\text{In}, \mathbf{D}_a; \mathbf{D}_b)$ outcome in the presence of a social tie, inequity aversion instead predicts that Alice will play (Out, \cdot) , no matter whether she is and/or expects

	C_b	D_b		C_b	D_b		C_b	D_b
C_a	(5,-25)	(0,0)	C_a	(5,5)	(0,0)	C_a	(5,5)	(5,15)
D_a	(0,0)	(-5,15)	D_a	(0,0)	(15,15)	D_a	(15,5)	(15,15)
a Inequity aversion model			b Fairness model			c Social tie model		

Fig. 7 Transformed social tie games without outside option

Bob to be inequity averse. However, our model of ties cannot be easily distinguished from the model of fairness presented in “[Theory of fairness](#)”: both theories may indeed predict the same outcome ($In, D_a; D_b$) can appear to be a unique social welfare equilibrium. In order to discriminate between those models, one may then consider a version of the ST game where the outside option is removed (that is, without the possibility for Alice to choose between *In* and *Out* before the coordination game): this simply corresponds to playing the coordination game alone. Figure 7 illustrates it through examples of the corresponding transformed payoff matrix when applying social preferences and the model of social ties to the reduced game.

According to Fig. 7a, both Alice and Bob are extremely inequity averse agents (i.e., $\alpha_a = \alpha_b = \beta_a = \beta_b = 1$), which leads them to miscoordinate by playing the (C_a, D_b) solution. Similarly, considering extremely fair agents (i.e., $\lambda = \delta = 1$ for Alice and Bob) as in Fig. 7b shows that both players cannot be expected to always coordinate: as the resulting game (which thus corresponds to a version of the Hi-Lo matching game) yields two different social welfare equilibria, both strategies can become rational for both players. However, considering our model of social ties (with $k_{ab} = k_{ba} = 1$), as in Fig. 7c, shows that both players should still coordinate on the (D_a, D_b) outcome, which corresponds to the unique Nash equilibrium of the resulting game. As the main result of this analysis, our model of ties clearly predicts that the outside option is irrelevant in the presence of a social tie between Alice and Bob. The players’ behavior remains the same independently of the presence of this outside option. In addition to the stability in the agents’ behavior it allows for, our model of ties appears to be more realistic than Charness and Rabin’s theory of Fairness. In fact, as shown through “[Theory of fairness](#)”, convergence towards a unique social welfare equilibrium requires a high level of fairness, along with some FI reasoning, and a propensity to help the worst-off person over maximizing the group payoff. Assuming that human beings have bounded computational resources, our model of social ties, which clearly relies on some low level of strategic reasoning, seems definitely more adequate to solve the sort of dilemma introduced by the ST game.

Furthermore, one should note that both our model of social ties and the concept of team reasoning happen to make the same predictions in the context of the ST game, no matter whether the outside option is present or not. Such an observation therefore suggests the need for investigating other relevant game theoretic situations in future work, as a means to disentangle predictions from those theories (see, e.g., the game in Table 5 from “[A theory of how to model social ties](#)”).

Conclusions

In this paper, we have proposed a game that appears to have very nice properties to investigate the behavioral effects of social ties. Indeed it creates a dilemma between maximizing self-interest and maximizing social welfare. It differs however from existing economic games from the experimental economic literature that elicit similar properties, such as the trust game, the ultimatum game, and the dictator game. Those games indeed provide situations where people’s decision may be influenced by some psychological factors (e.g., disappointment, regret, and guilt) (Geanakoplos et al. 1989). While investigating the impact of social ties on social emotions clearly

represents an interesting research orientation for future work, it is not the motivation of this paper: the strategic structure of the ST game introduced in “A social ties game” can hardly allow for such emotional reasoning. Moreover, a clear advantage of the ST game is that it is well suited to evaluate the very plausible theory of team reasoning in the context of social ties: the stronger the tie between individuals, the more they may act as members of the same group.

However, as this work is purely theoretical, it clearly suggests some further experimental analysis. The next stage of this study therefore consists of testing and evaluating the main hypotheses made in the previous section. To do so, we intend to conduct experimental sessions where people will be asked to interact (1) with some perfect strangers, and (2) with some socially connected individuals (e.g., friends, classmates, teammates, etc) in the context of our ST game in extensive form.

References

- Abrams, D., & Hogg, M. (2006). Comments on the motivational status of self-esteem in social identity and inter-group discrimination. *European Journal of Social Psychology*, 18(4), 317–334.
- Alger, I., & Weibull, J. (2012). *Homo moralis-preference evolution under incomplete information and assortative matching*. Tech. rep., Toulouse School of Economics (TSE).
- Bacharach, M. (1997). Weequilibria: a variable frame theory of cooperation. Unpublished paper presented at the seminar on co-operative reasoning, St. John's College, Oxford.
- Bacharach, M. (1999). Interactive team reasoning: a contribution to the theory of cooperation. *Research in Economics*, 23, 117–147.
- Bacharach, M. (2006). *Beyond individual choice: teams and frames in game theory*. Oxford: Princeton University Press.
- Bacharach, M., & Bernasconi, M. (1997). The variable frame theory of focal points: an experimental study. *Games and Economic Behavior*, 19(1), 1–45.
- Balkenborg, D. (1994). *An experiment on forward versus backward induction*. Rheinische Friedrich-Wilhelms-Universität Bonn, Bonn.
- Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior*, 10(1), 122–142.
- Binmore, K. (2005). *Natural justice*. New York: Oxford University Press.
- Binmore, K., & Samuelson, L. (1999). Evolutionary drift and equilibrium selection. *The Review of Economic Studies*, 66(2), 363.
- Bolton, G. E., & Ockenfels, A. (2000). A theory of equity, reciprocity and competition. *American Economic Review*, 100, 166–193.
- Brandts, J., Holt, C. (1989). *Forward induction: Experimental evidence from two-stage games with complete information*. Departament d'Economia i d'Història Econòmica, Universitat Autònoma de Barcelona.
- Brandts, J., & Holt, C. (1995). Limitations of dominance and forward induction: experimental evidence. *Economics Letters*, 49(4), 391–395.
- Brandts, J., Cabrales, A., Charness, G. (2003). *Forward induction and the excess capacity puzzle: an experimental investigation*. Paper provided by Department of Economics and Business, Universitat Pompeu Fabra.
- Cachon, G., & Camerer, C. (1996). Loss-avoidance and forward induction in experimental coordination games. *Quarterly Journal of Economics*, 111(1), 165.
- Caminati, M., Innocenti, A., & Ricciuti, R. (2006). Drift effect under timing without observability: experimental evidence. *Journal of Economic Behavior and Organization*, 61(3), 393–414.
- Charness, G. B., & Rabin, M. (2002). Understanding social preferences with simple tests. *Quarterly Journal of Economics*, 117, 817–869.
- Colman, A. M., Pulford, B. N., & Rose, J. (2008). Collective rationality in interactive decisions: evidence for team reasoning. *Acta Psychologica*, 128(2), 387–397.
- Cooper, R., De Jong, D., Forsythe, R., & Ross, T. (1992). Forward induction in coordination games. *Economics Letters*, 40(2), 167–172.

- Cooper, R., DeJong, D., Forsythe, R., Ross, T. (1993). Forward induction in the battle-of-the-sexes games. *The American Economic Review*, 83:1303–1316.
- Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, 114, 817–868.
- Frost, J., Chance, Z., Norton, M., & Ariely, D. (2008). People are experience goods: improving online dating with virtual dates. *Journal of Interactive Marketing*, 22(1), 51–61.
- Geanakoplos, J., Pearce, D., & Stacchetti, E. (1989). Psychological games and sequential rationality. *Games and Economic Behavior*, 1(1), 60–79.
- Gigerenzer, G., & Selten, R. (2001). *The adaptive toolbox. Bounded rationality: the adaptive toolbox* (pp. 37–50). Cambridge: MIT Press
- Gilbert, M. (1989). *On social facts*. London: Routledge.
- Gold, N. (2012). Team reasoning, framing and cooperation. In: S. Okasha, K. Binmore (Eds.), *Evolution and rationality: decision, cooperation and strategic behaviour* (pp. 185–212). New York: Cambridge University Press.
- Gold, N., & Sugden, R. (2007). *Theories of team agency*. In: Peter, F., Schmid H.B. (Ed) *Rationality and commitment*. Oxford: oxford University Press
- Güth, W., Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior and Organization*, 3(4), 367–388.
- Hakli, R., Miller, K., & Tuomela, R. (2010). Two kinds of we-reasoning. *Economics and Philosophy*, 26(3), 291.
- Hitsch, G., Hortaçsu, A., & Ariely, D. (2010). Matching and sorting in online dating. *The American Economic Review*, 100(1), 130–163.
- Hogg, M. (2000). Subjective uncertainty reduction through self-categorization: a motivational theory of social identity processes. *European Review of Social Psychology*, 11(1), 223–255.
- Hogg, M. (2002). Social identity. In M. Leary, J. Tangney (Eds.), *Handbook of self and identity* (pp. 462–479). New York: The Guilford Press.
- Huck, S., & Muller, W. (2005). Burning money and (pseudo) first-mover advantages: an experimental study on forward induction. *Games and Economic Behavior*, 51(1), 109–127.
- Lorini, E. (2011). From self-regarding to other-regarding agents in strategic games: a logical analysis. *Journal of Applied Non-Classical Logics*, 21(3–4), 443–475.
- Luhtanen, R., & Crocker, J. (1992). A collective self-esteem scale: self-evaluation of one's social identity. *Personality and Social Psychology Bulletin*, 18(3), 302–318.
- Margolis, H. (1982). *Selfishness, altruism, and rationality: a theory of social choice*. Chicago: University of Chicago Press.
- Rabin, M. (1993a). Incorporating fairness into game theory and economics. *American Economic Review*, 83(5), 1281–1302.
- Rabin, M. (1993). Incorporating fairness into game theory and economics. *The American Economic Review* 83: 1281–1302.
- Rawls, J. (1971). *A theory of justice*. Cambridge: Harvard University Press.
- Regan, D. (1980). *Utilitarianism and cooperation*. Oxford: Clarendon.
- Shafir, E., & Tversky, A. (1992). Thinking through uncertainty: nonconsequential reasoning and choice. *Cognitive Psychology*, 24(4), 449–474.
- Shahriar, Q. (2009). *Forward induction works! an experimental study to test the robustness and the power*. Working Papers.
- Smerilli, A. (2010). We-thinking and vacillation between frames: filling a gap in Bacharach's theory. *Theory and Decision* 1–22.
- Sugden, R. (2000). Team preferences. *Economics and Philosophy*, 16, 175–204.
- Sugden, R. (2003). The logic of team reasoning. *Philosophical Explorations*, 6(3), 165–181.
- Tajfel, H. (1970). Experiments in intergroup discrimination. *Scientific American*, 223(5), 96–102.
- Tajfel, H., & Turner, J. (1979). An integrative theory of intergroup conflict. *The social psychology of intergroup relations*. 33, 47.
- Tajfel, H., Billig, M., Bundy, R., & Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology*, 1(2), 149–178.
- Tuomela, R. (2010). *The philosophy of sociality: the shared point of view*. New York: Oxford University Press.
- Van Huyck John, B., Battalio Raymond, C., & Beil Richard, O. (1993). Asset markets as an equilibrium selection mechanism: coordination failure, game form auctions, and tacit communication. *Games and Economic Behavior*, 5(3), 485–504.